

Amendments to the Specification:

Please replace paragraph [0026] beginning at page 8, line 22, with the following:

--[0026] A consensus sequence may be represented by a formula, comprising both fixed and variable bases. Thus, the consensus sequence for the signature motif for the family of nuclear hormone receptors mentioned above and shown in Example 1 is:

[(A/T/G) (C/T) (A/G/T/C)] [TG (T/C)] [(A/G) (A/G) (A/C/G/T)]
[(A/G) (C/G) (A/C/G/T)] [TG (T/C)] [(A/T) (A/C/G) (A/C/G)]
[(A/G) (C/G/T) (A/C/G/T)] (SEQ ID NO:281)

As can be seen, among the variable bases, some may be fully random, *i.e.*, they may be any of the four possible bases, A, C, G or T. Others may be partially random, *i.e.*, they may comprise only two or only three predetermined bases of the four possible bases. Generally, in determining a consensus sequence, all possible codon variations for a given amino acid will be taken into account; however, for various reasons, including the need to limit the complexity (*i.e.* size) of the siRNA library, the consensus sequence may be restricted to include only the specific codons known to code for the amino acids comprising the known members of the protein family.--

Please replace paragraph [0040] (and Table 1) beginning at page 14, line 8, with the following:

--[0040] An additional and effective way to reduce the complexity of a library is to divide the members of a protein family of interest into two or more sub-sets, each sub-set comprising members having a variant of the signature sequence, each such variant comprising a relatively high number of amino acids that are constant for all members of the sub-set. The effect of such division can be seen clearly with reference to Example 2 and Table 1 below, which shows the

division into seven sub-sets of the 89 known members of the family of tyrosine kinases. Each of sub-sets 1 and 4-7 have a different variant of the signature motif, but all five comprise seven amino acids that are constant for all members of the respective sub-set. Sub-set 3 has a variant signature sequence in which only one of the seven amino acids is not constant for all members of the sub-set; and only sub-set 2 has a variant signature motif in which three of the amino acids are not constant for all members.

Table 1

Variant	Signature Motif	SEQ ID NO:	No. of Known Members	Complexity
1	H R D L K S S	282	3	4
2	H R N/D L/V/I A A/V R	283	3	2,304
3	H R D L R A/S A	284	8	10,368
4	H R/K D L A T R	285	9	2,592
5	H R D L A A R	286	61	8,192
6	H K D L A A R	287	3	576
7	H R D I A A R	288	2	32
Total			89	24,068

--

Please replace paragraph [0050] beginning at page 18, line 2, with the following:

--[0050] The symbols for amino acids used in the examples are as follows:

A	Alanine	M	Methionine
C	Cysteine	N	Asparagine
D	Aspartic acid	P	Proline
E	Glutamic acid	Q	Glutamine
F	Phenylalanine	R	Arginine
G	Glycine	S	Serine
H	Histidine	T	Threonine
I	Isoleucine	V	Valine
K	Lysine	W	Tryptophan
L	Leucine	Y	Tyrosine

Example 1

Family of human nuclear hormone receptors (ZnF_C4 domain) - 45 members

In this example, a single signature motif was designed based on the zinc finger domain present in all 45 known members of the nuclear hormone receptor family. A short segment of the zinc finger domain present in each of the 45 known family members is shown below. The consensus sequence was "reverse translated" utilizing only those codons that encode the signature motif region of known members of the family. Using a full 21-nucleotide consensus sequence to construct the siRNA library, the complexity would be 10,616,832. By reducing the length of the consensus sequence to 19 nucleotides, the complexity is reduced to 884,736. SiRNAs as short as 19 nucleotides are highly efficient at reducing their cognate mRNA levels [Czauderna, F. *et al.*, *Nucl. Acids Res.* **31**: 2705-2716 (2003)], therefore, reducing the length of the consensus sequence will have little, if any, effect on the degree of silencing produced by members of the library.

tataatgcactgacctgtgaggggtgtaaagggtttcttcaggaga (SEQ ID NO:1)
Y N A L T C E G C K G F F R R (SEQ ID NO:2)

tacggcgtgcgcacctgtgagggctgcaaaggcttctttaagcgc (SEQ ID NO:3)
Y G V R T C E G C K G F F K R (SEQ ID NO:4)

tacggcgtgcgcacctgtgagggctgcaaaggcttctttaagcgc (SEQ ID NO:5)
Y G V R T C E G C K G F F K R (SEQ ID NO:6)

tacggcgtgcgcaacctgcgagggctgcaaagggtttttcaagaga (SEQ ID NO:7)
Y G V R T C E G C K G F F K R (SEQ ID NO:8)

tatggtgtccgcacatgtgagggctgcaaaggcttcttcaagcgc (SEQ ID NO:9)
Y G V R T C E G C K G F F K R (SEQ ID NO:10)

tatggagcagtaacttgtgaaggctgcaaaggattttttaaaaga (SEQ ID NO:11)
Y G A V T C E G C K G F F K R (SEQ ID NO:12)

tacgggggttatcacctgtgaggggtgcaaaggcttcttccgccgg (SEQ ID NO:13)
Y G V I T C E G C K G F F R R (SEQ ID NO:14)

tacggagtcacacatgtgaaggctgcaaaggattctttaggagg (SEQ ID NO:15)
Y G V I T C E G C K G F F R R (SEQ ID NO:16)

tatggtgtcattacatgtgaaggctgcaaaggctttttcaggaga (SEQ ID NO:17)
Y G V I T C E G C K G F F R R (SEQ ID NO:18)

tatggagtgtacagctgaggggtgcaaaggcttcttcaagcgg (SEQ ID NO:19)
Y G V Y S C E G C K G F F K R (SEQ ID NO:20)

tacgggggttacagctgtgaggggtgcaaaggcttcttcaaacgc (SEQ ID NO:21)
Y G V Y S C E G C K G F F K R (SEQ ID NO:22)

tacgggggtatacagttgtgaaggctgcaaagggttcttcaagagg (SEQ ID NO:23)
Y G V Y S C E G C K G F F K R (SEQ ID NO:24)

tacaacgtgctcagctgcaaggctgcaaaggcttcttccggcgc (SEQ ID NO:25)
Y N V L S C E G C K G F F R R (SEQ ID NO:26)

tacaatgttctgagctgaggggtgcaaaggattcttccgccgc (SEQ ID NO:27)
Y N V L S C E G C K G F F R R (SEQ ID NO:28)

tatgggatcatctcctgtgaggggtgcaaagggtttttcaagcgg (SEQ ID NO:29)
Y G I I S C E G C K G F F K R (SEQ ID NO:30)

tatgggggtcagctcttgtgaaggctgcaagggttctttcgccga (SEQ ID NO:31)

Y G V S S C E G C K G F F R R (SEQ ID NO:32)

tatgggggtcagctcttgtgaaggctgcaagggttctttcgccga (SEQ ID NO:33)

Y G V S S C E G C K G F F R R (SEQ ID NO:34)

tatgggggtcagctcttgtgaagggtgcaaagggttcttcaaaagg (SEQ ID NO:35)

Y G A V S C E G C K G F F K R (SEQ ID NO:36)

tacgggtgtcttcacctgcgagggctgcaagagctttttcaagcga (SEQ ID NO:37)

Y G V F T C E G C K S F F K R (SEQ ID NO:38)

tacggccagttcacgtgcgagggctgcaagagcttcttcaagcgc (SEQ ID NO:39)

Y G Q F T C E G C K S F F K R (SEQ ID NO:40)

tacgggggtctacgcctgcgacggctgctcaggttttttcaaacgg (SEQ ID NO:41)

Y G V Y A C D G C S G F F K R (SEQ ID NO:42)

tatggcatctatgcctgcaacgggtgcagcggcttcttcaagagg (SEQ ID NO:43)

Y G I Y A C N G C S G F F K R (SEQ ID NO:44)

tatggggcatccacctgtgatgggtgcaagggttcttccagacgc (SEQ ID NO:45)

Y G A S T C D G C K G F F R R (SEQ ID NO:46)

tacgggtgcctcgagctgtgacgggtgcaagggttcttccggagg (SEQ ID NO:47)

Y G A S S C D G C K G F F R R (SEQ ID NO:48)

tatgggggtcagcgcctgtgaggggtgcaagggttcttccgcgcg (SEQ ID NO:49)

Y G V S A C E G C K G F F R R (SEQ ID NO:50)

tatgggggtcagcgcctgtgaggggatgtaagggttcttccgcaga (SEQ ID NO:51)

Y G V S A C E G C K G F F R R (SEQ ID NO:52)

tacgggtgtgcacgcctgcgaggggtgcaagggttcttccgtcgg (SEQ ID NO:53)

Y G V H A C E G C K G F F R R (SEQ ID NO:54)

tatggagttcatgcttgcgaagggtgtaagggttcttccggaga (SEQ ID NO:55)

Y G V H A C E G C K G F F R R (SEQ ID NO:56)

tacgggtgttcatgcatgtgaggggtgcaagggttcttccgtcgt (SEQ ID NO:57)

Y G V H A C E G C K G F F R R (SEQ ID NO:58)

tacggagtcacgcgtgtgaaggctgcaagggttctttcggcga (SEQ ID NO:59)

Y G V H A C E G C K G F F R R (SEQ ID NO:60)

tatggagttcatgcttgtgaaggatgcaagggttcttccggaga (SEQ ID NO:61)

Y G V H A C E G C K G F F R R (SEQ ID NO:62)

ttcaatgtcatgacatgtgaaggatgcaagggtttttcaggagg (SEQ ID NO:63)

F N V M T C E G C K G F F R R (SEQ ID NO:64)

tttaatgcgctgacttgtgagggctgcaagggttcttccaggaga (SEQ ID NO:65)

F N A L T C E G C K G F F R R (SEQ ID NO:66)

taccgctgtatcacgtgtgaaggctgcaagggttcttttagaaga (SEQ ID NO:67)

Y R C I T C E G C K G F F R R (SEQ ID NO:68)

taccgctgtatcacttgtgagggctgcaagggttctttcgccgc (SEQ ID NO:69)

Y R C I T C E G C K G F F R R (SEQ ID NO:70)

tacggactgctcacgtgtgagagctgcaagggttcttcaagcgc (SEQ ID NO:71)

Y G L L T C E S C K G F F K R (SEQ ID NO:72)

tatgggctcctcacctgtgaaagctgcaagggttttttaagcga (SEQ ID NO:73)

Y G L L T C E S C K G F F K R (SEQ ID NO:74)

tatggggtagtcacctgtggcagctgcaaagttttcttcaaaaga (SEQ ID NO:75)

Y G V V T C G S C K V F F K R (SEQ ID NO:76)

tatggagctctcacatgtggaagctgcaagggtcttcttcaaaaga (SEQ ID NO:77)

Y G A L T C G S C K V F F K R (SEQ ID NO:78)

tatggtgtccttacctgtgggagctgtaagggtcttctttaaagagg (SEQ ID NO:79)

Y G V L T C G S C K V F F K R (SEQ ID NO:80)

tatggagtccttaacttgtggaagctgtaaagttttcttcaaaaga (SEQ ID NO:81)

Y G V L T C G S C K V F F K R (SEQ ID NO:82)

tacggcgtggcctcctgcgaggcttgaaggccttcttcaagagg (SEQ ID NO:83)

Y G V A S C E A C K A F F K R (SEQ ID NO:84)

tatggtgtggcatcctgtgaggcctgcaaagccttcttcaagagg (SEQ ID NO:85)

Y G V A S C E A C K A F F K R (SEQ ID NO:86)

tatggagtctggtcctgtgagggctgcaaggccttcttcaagaga (SEQ ID NO:87)
Y G V W S C E G C K A F F K R (SEQ ID NO:88)

tatggagtctggtcgtgtgaaggatgtaaggccttttttaaaga (SEQ ID NO:89)
Y G V W S C E G C K A F F K R (SEQ ID NO:90)

Signature Motif:

(T/S/A)-(D/E/G/N)-(G/S/A)-(C)-(K/S)-(A/G/S/V)

Consensus sequence (21 nt) (SEQ ID NO:281):

(A/T/G) (C/T) (A/G/T/C) TG (T/C) (A/G) (A/G) (A/C/G/T)
T/S/A C D/E/G/N

(A/G) (C/G) (A/C/G/T) TG (T/C) (A/T) (A/C/G) (A/C/G)
G/S/A C K/S

(A/G) (C/G/T) (A/C/G/T)
G/S/V/A

Complexity: $2^9 \times 3^4 \times 4^4 = 512 \times 81 \times 256 = 10,616,832$ members

Consensus sequence (19 nt) (SEQ ID NO:289):

(A/T/G) (C/T) (A/G/T/C) TG (T/C) (A/G) (A/G) (A/C/G/T)
T/S/A C D/E/G/N

(A/G) (C/G) (A/C/G/T) TG (T/C) (A/T) (A/C/G) (A/C/G)
G/S/A C K/S

(A/G) --
G/S/V/A

Complexity: $2^9 \times 3^3 \times 4^3 = 512 \times 27 \times 64 = 884,736$ members

Example 2

Family of tyrosine kinases – 89 members

This example shows the identification of seven variants of a portion of the catalytic domain of the family of tyrosine kinases. As shown in Table 1 above, these may then be used for the production of library of siRNAs targeting this domain having a reduced complexity of 24,068 unique members.

Variant 1: 3 members

gttcccatcatccaccgacgacctaagtccagcaacatattgatcctc (SEQ ID NO:91)
V P I I H R D L K S S N I L I L (SEQ ID NO:92)

gtgcccatacctgcaccgggacctaagtccagcaacatatttgctactt (SEQ ID NO:93)
V P I L H R D L K S S N I L L L (SEQ ID NO:94)

gtgcccatacctgcaccgggacctaagtccagcaacatatttgctactt (SEQ ID NO:95)
V P I L H R D L K S S N I L L L (SEQ ID NO:96)

Signature Motif (SEQ ID NO:282): H R D L K S S

Consensus Sequence (SEQ ID NO:290):

<u>CAC</u>	<u>CG (C/G)</u>	<u>GAC</u>	<u>CT (C/T)</u>	<u>AAG</u>	<u>TCC</u>	<u>AGC</u>
H	R	D	L	K	S	S

Complexity: $2^2 = 4$ members

Variant 2: 3 members

catggtatggtgcatagaaacctggctgcccgaacgtgctactcaag (SEQ ID NO:97)
H G M V H R N L A A R N V L L K (SEQ ID NO:98)

aagaattgcatccaccgggacgtggcagcgcgtaacgtgctgttgacc (SEQ ID NO:99)
K N C I H R D V A A R N V L L T (SEQ ID NO:100)

atcaactgcgtgcacagggacattgctgtccggaacatcctggtggcc (SEQ ID NO:101)
I N C V H R D I A V R N I L V A (SEQ ID NO:102)

Signature Motif (SEQ ID NO:283): H R D/N I/V/L A A/V R

Consensus Sequence (SEQ ID NO:291):

<u>CA (T/C)</u>	<u>(C/A) G (G/A)</u>	<u>(G/A) AC</u>	<u>(A/C/G) T (T/G)</u>	<u>GC (T/A)</u>	<u>G (T/C) (C/G)</u>
H	R	D/N)	(I/V/L)	A	(A/V)
<u>CG (A/T/G)</u>					
R					

Complexity: $2^8 \times 3^2 = 256 \times 9 = 2,304$ members

Variant 3: 8 members

atgaactacgtccaccgggaccttcgtgcagccaacatcctggtggga	(SEQ ID NO:103)
M N Y V <u>H R D L R A A</u> N I L V G	(SEQ ID NO:104)

atgaactatattcaccgagatcttcgggctgctaattctttagga	(SEQ ID NO:105)
M N Y I <u>H R D L R A A</u> N I L V G	(SEQ ID NO:106)

atgaattatatccatagagatctgcgatcagcaaacattctagtgggg	(SEQ ID NO:107)
M N Y I <u>H R D L R S A</u> N I L V G	(SEQ ID NO:108)

atgaactacattcaccgcgacctgagggcagccaacatcctggttggg	(SEQ ID NO:109)
M N Y I <u>H R D L R A A</u> N I L V G	(SEQ ID NO:110)

aagaattccatccaccgcgacctgcgggcggccaacatcctggtgtct	(SEQ ID NO:111)
M N S I <u>H R D L R A A</u> N I L V S	(SEQ ID NO:112)

aggaactacatccaccgagacctccgagctgccaacatcttggtctt	(SEQ ID NO:113)
R N Y I <u>H R D L R A A</u> N I L V S	(SEQ ID NO:114)

aagaactacattcaccgggacctgcgagcagctaattgttctggtctcc	(SEQ ID NO:115)
K N Y I <u>H R D L R A A</u> N V L V S	(SEQ ID NO:116)

cggaattatattcatcgtgaccttcgggctgccaacattctggtgtct	(SEQ ID NO:117)
R N Y I <u>H R D L R A A</u> N I L V S	(SEQ ID NO:118)

Signature Motif (SEQ ID NO:284): H R D L R A/S A

Consensus Sequence (SEQ ID NO:292):

CA (C/T) (C/A) G (A/C/G/T) GA (C/T) CT (C/G/T) (A/C) G (A/G/T)
H R D L R

(G/T) C (A/G/T) GC (A/C/T)
(A/S) A

Complexity: $2^5 \times 3^4 \times 4 = 32 \times 81 \times 4 = 10,368$ members

Variant 4: 9 Members

ctgcattttgtgcaccgggacctggccacacgcaactgtctagtggg (SEQ ID NO:119)
L H F V H R D L A T R N C L V G (SEQ ID NO:120)

ctcaactttgtacatcgggacctggccacgcggaactgcctagtgtggg (SEQ ID NO:121)
L N F V H R D L A T R N C L V G (SEQ ID NO:122)

cttaattttgttcaccgagatctggccacacgaaactgtttagtgggt (SEQ ID NO:123)
L N F V H R D L A T R N C L V G (SEQ ID NO:124)

cgcgggctggtgcaccgagacctcgctacgcgcaacctactgctggcg (SEQ ID NO:125)
R G L V H R D L A T R N L L L A (SEQ ID NO:126)

aaaaggatatatccacagggatctggcaacgagaaatatattggtggag (SEQ ID NO:127)
K R Y I H R D L A T R N I L V E (SEQ ID NO:128)

cagcactttgtgcaccgagacctggccaccaggaactgcctgggttggg (SEQ ID NO:129)
Q H F V H R D L A T R N C L V G (SEQ ID NO:130)

cagcacttcgtgcaccgcgatttggccaccaggaactgcctgggtcggg (SEQ ID NO:131)
Q H F V H R D L A T R N C L V G (SEQ ID NO:132)

caccacgtgggttcacaaggacctggccaccgcgaatgtgctagtgtac (SEQ ID NO:133)
H H V V H K D L A T R N V L V Y (SEQ ID NO:134)

cgtaagtttgttcaccgagatttagccaccaggaactgcctgggtgggc (SEQ ID NO:135)
R K F V H R D L A T R N C L V G (SEQ ID NO:136)

Signature Motif (SEQ ID NO:285): H R/K D L A T R

Consensus Sequence (SEQ ID NO:293):

CAC (A/C) (A/G) (A/C/G) GA (C/T) (C/T) T (A/C/G) GC (A/C/T)
H R/K D L A

AC (A/C/G) (A/C) G (A/C/G)
T R

Complexity: $2^5 \times 3^5 = 32 \times 81 = 2,592$ members

Variant 5: 61 members

aagaagcttgtgcaccgcgacctggccgcccgaacatcctggtctca (SEQ ID NO:137)
K K L V H R D L A A R N I L V S (SEQ ID NO:138)

aagaagcttgtgcaccgggacctagccgcccgaacatcctggtctca (SEQ ID NO:139)
K K L V H R D L A A R N I L V S (SEQ ID NO:140)

aacaatttcgtgcatcgagacctggctgcccgaatgtgctggtgtct (SEQ ID NO:141)
N N F V H R D L A A R N V L V S (SEQ ID NO:142)

cacgactacatccaccgagacctagccgcgcgaacgtgctgctggac (SEQ ID NO:143)
H D Y I H R D L A A R N V L L D (SEQ ID NO:144)

cggcaatacgttcaccgggacttggcagcaagaaatgtccttggtgag (SEQ ID NO:145)
R Q Y V H R D L A A R N V L V E (SEQ ID NO:146)

cgctcgcttggtgcaccgcgacctggcagccaggaacgtactggtgaaa (SEQ ID NO:147)
R R L V H R D L A A R N V L V K (SEQ ID NO:148)

cggaacttcacccaccgagacctggctgctcggaattgcatgctggca (SEQ ID NO:149)
R N F I H R D L A A R N C M L A (SEQ ID NO:150)

aagaagtgcatacaccgagacctggcagccaggaatgtcctggtgaca (SEQ ID NO:151)
K K C I H R D L A A R N V L V T (SEQ ID NO:152)

caaaaatgtattcatcgagatttagcagccagaaatgttttggttaaca (SEQ ID NO:153)
Q K C I H R D L A A R N V L V T (SEQ ID NO:154)

cagaagtgcacccacaggacctggctgcccgaatgtgctggtgacc (SEQ ID NO:155)
Q K C I H R D L A A R N V L V T (SEQ ID NO:156)

cagaagtgtattcacagagacttggctgccagaaacgtcctggtgacc (SEQ ID NO:157)

Q K C I H R D L A A R N V L V T (SEQ ID NO:158)

cggaagtgtatccaccgggacctggctgcccgcaatgtgctggtgact (SEQ ID NO:159)

R K C I H R D L A A R N V L V T (SEQ ID NO:160)

atgaagctcgttcacatcgaggacttggcagccagaaacatcctggtagct (SEQ ID NO:161)

M K L V H R D L A A R N I L V A (SEQ ID NO:162)

agaaagtgcattcatcgggacctggcagcgagaaacattcttttatct (SEQ ID NO:163)

R K C I H R D L A A R N I L L S (SEQ ID NO:164)

cgaaagtgcattcacagagacctggctgctcggaacattctgctgtcg (SEQ ID NO:165)

R K C I H R D L A A R N I L L S (SEQ ID NO:166)

cgaaagtgtatccacagggacctggcggcacgaaatatcctcttatcg (SEQ ID NO:167)

R K C I H R D L A A R N I L L S (SEQ ID NO:168)

aagaactgcgtccacagagacctggcggctaggaacgtgctcatctgt (SEQ ID NO:169)

K N C V H R D L A A R N V L I C (SEQ ID NO:170)

aaaaattgtgtccaccgtgatctggctgctcgcaacgtcctcctggca (SEQ ID NO:171)

K N C V H R D L A A R N V L L A (SEQ ID NO:172)

aagaattgtattcacagagacttggcagccagaaatatcctccttact (SEQ ID NO:173)

K N C I H R D L A A R N I L L T (SEQ ID NO:174)

aagtcgtgtgttcacagagacctggccgcccaggaacgtgcttgtcacc (SEQ ID NO:175)

K S C V H R D L A A R N V L V T (SEQ ID NO:176)

aaacagttttattcacagggacctagctgccaggaacatttttagttggt (SEQ ID NO:177)

K Q F I H R D L A A R N I L V G (SEQ ID NO:178)

aagcagttcatccacagggacctggctgcccggaatgtgctggtcgga (SEQ ID NO:179)

K Q F I H R D L A A R N V L V G (SEQ ID NO:180)

aagcagttcatccacagggacctggctgcccggaatgtgctggtcgga (SEQ ID NO:181)

K Q F I H R D L A A R N V L V G (SEQ ID NO:182)

atgaactatgtgcaccgtgacctggctgcccgcaacatcctcgtaac (SEQ ID NO:183)

M N Y V H R D L A A R N I L V N (SEQ ID NO:184)

atgcatttcattcacagggatctggcagctagaaattgccttgtttcc (SEQ ID NO:185)

M H F I H R D L A A R N C L V S (SEQ ID NO:186)

aacaagtttgtgcaccgagatctagcagcccgcaactgcatgggtgtcc (SEQ ID NO:187)

N K F V H R D L A A R N C M V S (SEQ ID NO:188)

aataagttcgtccacagagaccttgctgcccgggaattgcatggtagcc (SEQ ID NO:189)

N K F V H R D L A A R N C M V A (SEQ ID NO:190)

aagaagtttgtgcatcgaggacctggcagcgagaaactgcatgggtcgcc (SEQ ID NO:191)

K K F V H R D L A A R N C M V A (SEQ ID NO:192)

aagagattcatacacgggacctggcggccaggaactgcatgctgaat (SEQ ID NO:193)

K R F I H R D L A A R N C M L N (SEQ ID NO:194)

atgaactatgttcaccgtgacctggctgcccgcaacatcctcgtcaac (SEQ ID NO:195)

M N Y V H R D L A A R N I L V N (SEQ ID NO:196)

atgaactatgtgcaccgcgacctggctgctcgcaacatccttgtcaac (SEQ ID NO:197)

M N Y V H R D L A A R N I L V N (SEQ ID NO:198)

atgaattatgtgcatcgaggacctggctgctaggaacattctgggtcaac (SEQ ID NO:199)

M N Y V H R D L A A R N I L V N (SEQ ID NO:200)

atgggctatgtgcatagagatcttgctgccagaaacatcttaatcaac (SEQ ID NO:201)

M G Y V H R D L A A R N I L I N (SEQ ID NO:202)

cagaagtttgtgcacagggacctggctgcgcggaactgcatgctggac (SEQ ID NO:203)

Q K F V H R D L A A R N C M L D (SEQ ID NO:204)

aaaaagtttgtccacagagaccttggtgcaagaaactgtatgctggat (SEQ ID NO:205)

K K F V H R D L A A R N C M L D (SEQ ID NO:206)

atgggctatgttcaccgagacctcgctgctcggaacatcttgatcaac (SEQ ID NO:207)

M G Y V H R D L A A R N I L I N (SEQ ID NO:208)

aggaattttcttcacgagatttagctgctcgaaactgcatgttgcca (SEQ ID NO:209)

R N F L H R D L A A R N C M L R (SEQ ID NO:210)

aaaaactgtatacacagggaccttgctgcaagaaactgcctggtaggt (SEQ ID NO:211)

K N C I H R D L A A R N C L V G (SEQ ID NO:212)

aagtgcctgcacccacgggacctggctgctcggaactgcctggtgaca (SEQ ID NO:213)

K C C I H R D L A A R N C L V T (SEQ ID NO:214)

atgagctatgtgcacgcgtgatctggccgcacggaacatcctggtgaac (SEQ ID NO:215)

M S Y V H R D L A A R N I L V N (SEQ ID NO:216)

atgagctacgtccaccgagacctggctgctcgcaacatcctagtcaac (SEQ ID NO:217)

M S Y V H R D L A A R N I L V N (SEQ ID NO:218)

atgggctatgttcaccgagacctcgctgctcggaacatcttgatcaac (SEQ ID NO:219)

M G Y V H R D L A A R N I L I N (SEQ ID NO:220)

atgggatatgttcacagggaccttgagctcgcaatattcttgtcaac (SEQ ID NO:221)

M G Y V H R D L A A R N I L V N (SEQ ID NO:222)

cgtcgcttggtgcaccgcgacctggcagccaggaacgtactggtgaaa (SEQ ID NO:223)

R R L V H R D L A A R N V L V K (SEQ ID NO:224)

gtgcggctcgtacacagggacttgccgctcggaacgtgctggtcaag (SEQ ID NO:225)

V R L V H R D L A A R N V L V K (SEQ ID NO:226)

agacgactcgttcacgcgggatttggcagcccgtaatgtcttagtgaaa (SEQ ID NO:227)

R R L V H R D L A A R N V L V K (SEQ ID NO:228)

aaaaacttcacccacagagatcttgctgcccgaaactgcctggtaggg (SEQ ID NO:229)

K N F I H R D L A A R N C L V G (SEQ ID NO:230)

aagcgctttattcacctgacctggctgcccgcaatctgctggttggt (SEQ ID NO:231)

K R F I H R D L A A R N L L L A (SEQ ID NO:232)

aagaactttgtgcaccgtgacctggcggcccgcaacgtcctgctggtt (SEQ ID NO:233)

K N F V H R D L A A R N V L L V (SEQ ID NO:234)

aagaactttgtgcaccgtgacctggcggcccgcaacgtcctgctggtt (SEQ ID NO:235)

K N F V H R D L A A R N V L L V (SEQ ID NO:236)

agcaattttgtgcacagagatctggctgcaagaaatgtggtgctagtt (SEQ ID NO:237)

S N F V H R D L A A R N V L L V (SEQ ID NO:238)

cagaattacatccaccgggacctggccgccaggaacatcctcgtcggg (SEQ ID NO:239)

Q N Y I H R D L A A R N I L V G (SEQ ID NO:240)

cagcgcgttgtgcaccgggacttggccgcccggaaacgtgctcgtggac (SEQ ID NO:241)

Q R V V H R D L A A R N V L V D (SEQ ID NO:242)

cggaactacattcacagagatctggctgccagaaatgtcctcgttggt (SEQ ID NO:243)

R N Y I H R D L A A R N V L V G (SEQ ID NO:244)

aagaatttcattccatagagatcttgcagctcgtaactgcctagtggga (SEQ ID NO:245)

K N F I H R D L A A R N C L V G (SEQ ID NO:246)

aacagcttcattccacagagatctggctgccagaaattgtctagtaagt (SEQ ID NO:247)

N S F I H R D L A A R N C L V S (SEQ ID NO:248)

aatggctatatattcatagggatttggcggcaaggaattgtttggtcagt (SEQ ID NO:249)

N G Y I H R D L A A R N C L V S (SEQ ID NO:250)

gcatgtgtcatccacagagacttggctgccagaaattgtttggtggga (SEQ ID NO:251)

A C V I H R D L A A R N C L V G (SEQ ID NO:252)

caccaattcatacaccgggacttggctgctcgtaactgcttggtggac (SEQ ID NO:253)

H Q F I H R D L A A R N C L V D (SEQ ID NO:254)

aagcagttccttcaccgagacctggcagctcgaaactgtttggtaaac (SEQ ID NO:255)

K Q F L H R D L A A R N C L V N (SEQ ID NO:256)

cacaattatgtccaccgggacctggctgccagaaacatcttggtgaat (SEQ ID NO:257)

H N Y V H R D L A A R N I L V N (SEQ ID NO:258)

Signature Motif (SEQ ID NO:286): H R D L A A R

Consensus Sequence (SEQ ID NO:294):

<u>CA (C/T)</u>	<u>(A/C) G (A/C/G/T)</u>	<u>GA (C/T)</u>	<u>(T/C) T (A/C/G/T)</u>	<u>GC (A/C/G/T)</u>
H	R	D	L	A

<u>GC (A/C/G/T)</u>	<u>(A/C) G (A/C/G/T)</u>
A	R

Complexity: $2^4 \times 4^5 = 16 \times 512 = 8,192$ members

Variant 6: 3 members

agggaagtcatccacaaagacctggctgccaggaactgtgtcattgat (SEQ ID NO:259)
R E V I H K D L A A R N C V I D (SEQ ID NO:260)

aaccgctttgtgcataaggacttggtgctgcgcgtaactgcctggtcagt (SEQ ID NO:261)
N R F V H K D L A A R N C L V S (SEQ ID NO:262)

cacttctttgtccacaaggaccttgcagctcgcaatatatttaaatcgga (SEQ ID NO:263)
H F F V H K D L A A R N I L I G (SEQ ID NO:264)

Signature Motif (SEQ ID NO:287): H K D L A A R

Consensus Sequence (SEQ ID NO:295):

CA (C/T) AA (A/G) GAC (C/T) T (G/T) GC (A/T) GC (C/G/T) A/C) G (C/G/T)
H K D L A A R

Complexity: $2^6 \times 3^2 = 64 \times 9 = 576$ members

Variant 7: 2 members

aatcacttcatccacagggatattgccgcccgggaactgcctgctgagc (SEQ ID NO:265)
N H F I H R D I A A R N C L L S (SEQ ID NO:266)

aaccacttcatccaccgagacattgctgccagaaactgcctcttgacc (SEQ ID NO:267)
N H F I H R D I A A R N C L L T (SEQ ID NO:268)

Signature Motif (SEQ ID NO:288): H R D I A A R

Consensus Sequence (SEQ ID NO:296):

CAC (A/C) G (A/G) GA (C/T) ATT GC (C/T) GCC (A/C) G (A/G)
H R D I A A R

Complexity: $2^5 = 32$ $2^6 = 64$ members

Example 3

Family of human nuclear hormone receptors (ZnF_C4 domain) - 45 members divided into 9 groups

In this example, the 45 known members of the nuclear hormone receptor family are divided into 9 subgroups. The same segment of the Zinc Finger_C4 domain described in Example 1 was used to design individual signature motifs and consensus sequences for each of the 9 subgroups. As in Example 1, the consensus sequence was "reverse translated" utilizing only those codons that encode the signature motif region of known members of the subgroup. Division of the family into subgroups dramatically reduces the complexity from 10,616,832 (see Example 1) to 1,664.

Variant 1: 9 members

tataatgcactgacctgtgaggggtgtaaagggtttcttcaggaga	(SEQ ID NO:1)
Y N A L <u>T C E G C K G</u> F F R R	(SEQ ID NO:2)
tacggcgtgcgcacctgtgagggctgcaaaggcttctttaagcgc	(SEQ ID NO:3)
Y G V R <u>T C E G C K G</u> F F K R	(SEQ ID NO:4)
tacggcgtgcgcacctgtgagggctgcaaaggcttctttaagcgc	(SEQ ID NO:5)
Y G V R <u>T C E G C K G</u> F F K R	(SEQ ID NO:6)
tacggcgtgcgaaacctgcgagggctgcaaaggctttttcaagaga	(SEQ ID NO:7)
Y G V R <u>T C E G C K G</u> F F K R	(SEQ ID NO:8)
tatggtgtccgcacatgtgagggctgcaaaggcttcttcaagcgc	(SEQ ID NO:9)
Y G V R <u>T C E G C K G</u> F F K R	(SEQ ID NO:10)
tatggagcagtaacttgtgaaggctgcaaaggattttttaaaaga	(SEQ ID NO:11)
Y G A V <u>T C E G C K G</u> F F K R	(SEQ ID NO:12)
tacggggttatcacctgtgaggggtgcaaaggcttcttccgcgg	(SEQ ID NO:13)
Y G V I <u>T C E G C K G</u> F F R R	(SEQ ID NO:14)
tacggagtcatcacatgtgaaggctgcaaaggattctttaggagg	(SEQ ID NO:15)
Y G V I <u>T C E G C K G</u> F F R R	(SEQ ID NO:16)

tatgggtgtcattacatgtgaaggctgcaagggttttttcaggaga (SEQ ID NO:17)

Y G V I T C E G C K G F F R R (SEQ ID NO:18)

Signature Motif (SEQ ID NO:297): T C E G C K G

Consensus Sequence (SEQ ID NO:298):

<u>A-C- (A/C/T)</u>	<u>T-G- (C/T)</u>	<u>G-A- (A/G)</u>	<u>G-G- (C/G)</u>	<u>T-G- (C/T)</u>
T	C	E	G	C

<u>A-A- (A/G)</u>	<u>G-G- (A/C/T)</u>
K	G

Complexity: $2^5 \times 3^2 = 32 \times 9 = 288$ **Variant 2: 9 members**

tatggagtgtacagctgcgaggggtgcaagggttcttcaagcgg (SEQ ID NO:19)

Y G V Y S C E G C K G F F K R (SEQ ID NO:20)

tacgggggtttacagctgtgaggggtgcaagggttcttcaaacgc (SEQ ID NO:21)

Y G V Y S C E G C K G F F K R (SEQ ID NO:22)

tacgggggtatacagttgtgaaggctgcaaagggttcttcaagagg (SEQ ID NO:23)

Y G V Y S C E G C K G F F K R (SEQ ID NO:24)

tacaacgtgctcagctgcgagggtgcaagggttcttccggcgc (SEQ ID NO:25)

Y N V L S C E G C K G F F R R (SEQ ID NO:26)

tacaatgttctgagctgcgaggggtgcaagggttcttccggcgc (SEQ ID NO:27)

Y N V L S C E G C K G F F R R (SEQ ID NO:28)

tatgggatcatctcctgtgaggggtgcaaagggtttttcaagcgg (SEQ ID NO:29)

Y G I I S C E G C K G F F K R (SEQ ID NO:30)

tatgggggtcagctcttgtgaagggtgcaagggttctttcgccga (SEQ ID NO:31)

Y G V S S C E G C K G F F R R (SEQ ID NO:32)

tatgggggtcagctcttgtgaaggctgcaagggttctttcgccga (SEQ ID NO:33)

Y G V S S C E G C K G F F R R (SEQ ID NO:34)

tatggggctgtcagttgtgaagggttcaaagggttcttcaaaagg (SEQ ID NO:35)

Y G A V S C E G C K G F F K R (SEQ ID NO:36)

Signature Motif (SEQ ID NO:299): S C E G C K G

Consensus Sequence (SEQ ID NO:300):

<u>(A/T) - (C/G) - (C/T)</u>	<u>T-G- (C/T)</u>	<u>G-A- (A/G)</u>	<u>G-G- (C/G/T)</u>	<u>T-G-C</u>
S	C	E	G	C

<u>A-A- (A/G)</u>	<u>G-G- (A/C/G/T)</u>
K	G

Complexity: $2^6 \times 3 \times 4 = 64 \times 3 \times 4 = 768$ **Variant 3:** 2 members

tacgggtgtcttcacctgcgagggctgcaagagctttttcaagcga (SEQ ID NO:37)

Y G V F T C E G C K S F F K R (SEQ ID NO:38)

tacggccagttcacgtgcgagggctgcaagagcttcttcaagcgc (SEQ ID NO:39)

Y G Q F T C E G C K S F F K R (SEQ ID NO:40)

Signature Motif (SEQ ID NO:301): T C E G C K S

Consensus Sequence (SEQ ID NO:302):

<u>A-C- (C/G)</u>	<u>T-G-C</u>	<u>G-A-G</u>	<u>G-G-C</u>	<u>T-G-C</u>	<u>A-A- (A/G)</u>	<u>A-G- (C/T)</u>
T	C	E	G	C	K	S

Complexity: $2^3 = 8$

Variant 4: 2 members

tacgggggtctacgcctgcgacgggtgctcaggttttttcaaacgg (SEQ ID NO:41)
Y G V Y A C D G C S G F F K R (SEQ ID NO:42)

tatggcatctatgcctgcaacgggtgcagcggcttcttcaagagg (SEQ ID NO:43)
Y G I Y A C N G C S G F F K R (SEQ ID NO:44)

Signature Motif (SEQ ID NO:303): A C D/N G C S G

Consensus Sequence (SEQ ID NO:304):

$\frac{G-C-C}{A}$ $\frac{T-G-C}{C}$ $\frac{(A/G)-A-C}{D/N}$ $\frac{G-G-C}{G}$ $\frac{T-G-C}{C}$ $\frac{(A/T)-(C/G)-(A/C)}{S}$

$\frac{G-G-(C/T)}{G}$

Complexity: $2^5 = 32$

Variant 5: 2 members

tatggggcatccacctgtgatgggtgcaagggtttcttcagacgc (SEQ ID NO:45)
Y G A S T C D G C K G F F R R (SEQ ID NO:46)

tacgggtgcctcgagctgtgacgggtgcaagggttcttccggagg (SEQ ID NO:47)
Y G A S S C D G C K G F F R R (SEQ ID NO:48)

Signature Motif (SEQ ID NO:305): T/S C D G C K G

Consensus Sequence (SEQ ID NO:306):

$\frac{A-(C/G)-C}{S/T}$ $\frac{T-G-T}{C}$ $\frac{G-A-(C/T)}{D}$ $\frac{G-G-(C/G)}{G}$ $\frac{T-G-C}{C}$ $\frac{A-A-G}{K}$

$\frac{G-G-(C/T)}{G}$

Complexity: $2^4 = 16$

Variant 6: 7 members

tatgggggtcagcgccctgtgaggggtgcaagggtttcttccgccgc	(SEQ ID NO:49)
Y G V S <u>A C E G C K G</u> F F R R	(SEQ ID NO:50)
tatgggggtcagcgccctgtgaggggatgtaagggtttttccgcaga	(SEQ ID NO:51)
Y G V S <u>A C E G C K G</u> F F R R	(SEQ ID NO:52)
tacgggtgtgcacgcctgcgaggggtgcaagggtttttccgtcgg	(SEQ ID NO:53)
Y G V H <u>A C E G C K G</u> F F R R	(SEQ ID NO:54)
tatggagttcatgcttgccaagggtgtaagggtttctttcggaga	(SEQ ID NO:55)
Y G V H <u>A C E G C K G</u> F F R R	(SEQ ID NO:56)
tacgggtgttcatgcatgtgaggggtgcaagggtttcttccgtcgt	(SEQ ID NO:57)
Y G V H <u>A C E G C K G</u> F F R R	(SEQ ID NO:58)
tacggagtcacgcgtgtgaagggtgcaagggtttctttcggcga	(SEQ ID NO:59)
Y G V H <u>A C E G C K G</u> F F R R	(SEQ ID NO:60)
tatggagttcatgcttgtgaaggatgcaagggtttcttccggaga	(SEQ ID NO:61)
Y G V H <u>A C E G C K G</u> F F R R	(SEQ ID NO:62)

Signature Motif (SEQ ID NO:307): A C E G C K G

Consensus Sequence (SEQ ID NO:308):

<u>G-C- (A/C/G/T)</u>	<u>T-G- (C/T)</u>	<u>G-A- (A/G)</u>	<u>G-G- (A/C/G)</u>	<u>T-G- (C/T)</u>
A	C	E	G	C
<u>A-A-G</u>	<u>G-G- (C/T)</u>			
K	G			

Complexity: $2^4 \times 3 \times 4 = 16 \times 12 = 192$

Variant 7: 6 members

ttcaatgtcatgacatgtgaaggatgcaagggtttttcaggagg	(SEQ ID NO:63)
F N V M <u>T C E G C K G</u> F F R R	(SEQ ID NO:64)

tttaatgcgctgacttgtgagggctgcaagggtttcttcaggaga (SEQ ID NO:65)
F N A L T C E G C K G F F R R (SEQ ID NO:66)

taccgctgtatcacgtgtgaaggctgcaagggtttctttagaaga (SEQ ID NO:67)
Y R C I T C E G C K G F F R R (SEQ ID NO:68)

taccgctgtatcacttgtgagggctgcaagggtttctttcgccgc (SEQ ID NO:69)
Y R C I T C E G C K G F F R R (SEQ ID NO:70)

tacggactgctcacgtgtgagagctgcaagggtttcttcaagcgc (SEQ ID NO:71)
Y G L L T C E S C K G F F K R (SEQ ID NO:72)

tatgggctcctcacctgtgaaagctgcaagggttttttaagcga (SEQ ID NO:73)
Y G L L T C E S C K G F F K R (SEQ ID NO:74)

Signature Motif (SEQ ID NO:309): T C E G/S C K G

Consensus Sequence (SEQ ID NO:310):

A-C- (A/C/G/T) T-G-T G-A- (A/G) (A/G) -G- (A/C) T-G-C A-A-G
T C E G/S C K

G-G (A/C/T)
G

Complexity: $2^3 \times 3 \times 4 = 8 \times 3 \times 4 = 96$

Variant 8: 4 members

tatggggtagtcacctgtggcagctgcaaagttttcttcaaaaga (SEQ ID NO:75)
Y G V V T C G S C K V F F K R (SEQ ID NO:76)

tatggagctctcacatgtggaagctgcaagggtttcttcaaaaga (SEQ ID NO:77)
Y G A L T C G S C K V F F K R (SEQ ID NO:78)

tatggtgtccttacctgtgggagctgtaagggtttctttaagagg (SEQ ID NO:79)
Y G V L T C G S C K V F F K R (SEQ ID NO:80)

tatggagtccttaacttgtggaagctgtaaagttttcttcaaaaga (SEQ ID NO:81)
Y G V L T C G S C K V F F K R (SEQ ID NO:82)

Signature Motif (SEQ ID NO:311): T C G S C K V

Consensus Sequence (SEQ ID NO:312):

$\frac{A-C-(A/C/T)}{T}$ $\frac{T-G-T}{C}$ $\frac{G-G-(A/C/G)}{G}$ $\frac{A-G-C}{S}$ $\frac{T-G-(C/T)}{C}$ $\frac{A-A-(A/G)}{K}$

$\frac{G-T-(C/T)}{V}$

Complexity: $2^3 \times 3^2 = 8 \times 9 = 72$

Variant 9: 4 members

tacggcgtggcctcctgcgaggcttgcaaggccttcttcaagagg (SEQ ID NO:83)
Y G V A S C E A C K A F F K R (SEQ ID NO:84)

tatgggtgtggcatcctgtgaggcctgcaaagccttcttcaagagg (SEQ ID NO:85)
Y G V A S C E A C K A F F K R (SEQ ID NO:86)

tatggagtctggtcctgtgagggctgcaaggccttcttcaagaga (SEQ ID NO:87)
Y G V W S C E G C K A F F K R (SEQ ID NO:88)

tatggagtctggtcgtgtgaaggatgtaaggccttttttaaaga (SEQ ID NO:89)
Y G V W S C E G C K A F F K R (SEQ ID NO:90)

Signature Motif (SEQ ID NO:313): S C E A/G C K A

Consensus Sequence (SEQ ID NO:314):

$\frac{T-C-(C/G)}{S}$ $\frac{T-G-(C/T)}{C}$ $\frac{G-A-(A/G)}{E}$ $\frac{G-(C/G)-(A/C/T)}{A/G}$ $\frac{T-G-(C/T)}{C}$

$\frac{A-A-(A/G)}{K}$ $\frac{G-C-C}{A}$

Complexity: $2^6 \times 3 = 64 \times 3 = 192$

Total Complexity of library: the sum of the complexities of subgroups 1-9 = 1,664.

The library is constructed from the following semi-randomized oligonucleotides:

Variant 1 (SEQ ID NO:269)

5' -pCCAGGACGACAAAAAGACHTGYGARGGSTGYAARGGHCTTTTAGGCTTTTCGG-3'

Variant 2 (SEQ ID NO:270)

5' -pCCAGGACGACAAAAAGWSYTGYGARGGBTGCAARGGNCTTTTAGGCTTTTCGG-3'

Variant 3 (SEQ ID NO:271)

5' -pCCAGGACGACAAAAAGACSTGCGAGGGCTGCAARAGYCTTTTAGGCTTTTCGG-3'

Variant 4 (SEQ ID NO:272)

5' -pCCAGGACGACAAAAAGCCTGCRACGGCTGCWSMGGYCTTTTAGGCTTTTCGG-3'

Variant 5 (SEQ ID NO:273)

5' -pCCAGGACGACAAAAAGASCTGTGAYGGSTGCAAGGGYCTTTTAGGCTTTTCGG-3'

Variant 6 (SEQ ID NO:274)

5' -pCCAGGACGACAAAAAGCNTGYGARGGVTGYAAGGGYCTTTTAGGCTTTTCGG-3'

Variant 7 (SEQ ID NO:275)

5' -pCCAGGACGACAAAAAGACNTGTGARRGMTGCAAGGGHCTTTTAGGCTTTTCGG-3'

Variant 8 (SEQ ID NO:276)

5' -pCCAGGACGACAAAAAGACHTGTGGVAGCTGYAARGTYCTTTTAGGCTTTTCGG-3'

Variant 9 (SEQ ID NO:277)

5' -pCCAGGACGACAAAAAGTCSTGYGARGSHTGYAARGCCTTTTAGGCTTTTCGG-3'

In the above, mixtures of nucleotides (wobbles) are denoted using the following standard nomenclature:

Table 2

Wobble	Nucleotides
B	C+G+T
D	A+G+T
H	A+C+T
K	G+T
M	A+C
N	A+C+G+T
R	A+G
S	C+G
V	A+C+G
W	A+T
Y	C+T

The semi-randomized oligonucleotides are resuspended in TE buffer and combined in direct proportion to their complexities to a final concentration of 0.92 μ M. One hundred eight pmol of the semi-randomized oligonucleotide mixture is combined with 21.6 pmol each of adapter oligonucleotides Univ-1(FseI) and Univ-2(AscI).

Univ-1(FseI): 5' -CTTTTGTGTCGTCTGGCCGG-3' (SEQ ID NO:278)

Univ-2(AscI): 5' -pCGCGCCGAAAAGCCTAAAAAG-3' (SEQ ID NO:279)

The oligonucleotides are annealed by heating to 70 °C for 5 minutes and slowly cooling to room temperature (~3 hours). The annealed oligonucleotides are ligated to 0.216 pmol of an FseI/AscI-digested vector bearing opposing human U6 and murine U6 promoters. Construction of this vector is described in U.S. Patent Application Serial Number 10/626,512. The nucleotide sequence of the human U6 and murine U6 promoters between the TATA box and the transcription start site was modified to contain FseI and AscI restriction sites, respectively, as indicated below:

Human U6/murine U6 Opposing Promoter Cassette

(FseI and AscI sites in lower case letters):

```
GGATCCAAGCTTAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCC
TTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTA
ATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAG
TAATAATTTCTTGGGTAGTTTGCAGTTTTTAAAATTATGTTTTTAAAATGGA
CTATCATATGCTTACCGTAACTTGAAAGTATTTTCGATTTCTTGGCTTTAT
ATATCggccggccTCGAggcgcgccATATTTATAGTCTCAAAACACACAA
TTACTTTACAGTTAGGGTGAGTTTCCTTTTGTGCTGTTTTTTAAAATAAT
AATTTAGTATTTGTATCTCTTATAGAAATCCAAGCCTATCATGTAAAATG
TAGCTAGTATTAAAAAGAACAGATTATCTGTCTTTTATCGCACATTAAGC
CTCTATAGTTACTAGGAAATATTATATGCAAATTAACCGGGGCAGGGGAG
TAGCCGAGCTTCTCCCAAGTCTGTGCGAGGGGGCCGGCGCGGGCCTAG
AGATGGCGGCGTCGGATCC (SEQ ID NO:280)
```

Ligation is performed overnight at 16 °C. One-fifth of the ligation reaction is used to transform electrocompetent bacteria (DH12S), resulting in 10^6 - 10^7 cfu/ μ g DNA.

The relatively low complexity (1,664) permits the delivery of the resulting library to the host cells by transient transfection in a 96-well format. The library is arrayed by picking ~4,000 individual colonies and inoculating 750 μ l/well of TB media (containing appropriate antibiotics) in 2-ml deep well 96-well plates (VWR). Following incubation for 20 hours, the cultures are pooled in groups of 10. DNA minipreps (Qiaprep Spin Miniprep Kits, Qiagen) are prepared from 1.5 ml of pooled bacterial culture. (The remainder of each culture is aliquotted and frozen for future use.) The purified DNA from each pool is quantitated using Rediplate 96 PicoGreen dsDNA Quantitation Kits (Molecular Probes). DNA from each pool is diluted to 100 ng/ μ l and stored in 96-well plates. Each well contains DNA encoding up to 10 unique siRNAs. Transfection of target cells is performed in a 96-well format using standard methods.--

Appl. No. 10/776,399

PATENT

Amdt. dated June 25, 2004

Reply to Notice to File Missing Parts of May 18, 2004

Please insert the accompanying paper copy of the Sequence Listing, page numbers 1 to 58, at the end of the application.